

Universal covers, color refinement, and two-variable counting logic: Lower bounds for the depth

Andreas Krebs* and Oleg Verbitsky†

Abstract

Given a connected graph G and its vertex x , let $U_x(G)$ denote the universal cover of G obtained by unfolding G into a tree starting from x . Let $T = T(n)$ be the minimum number such that, for graphs G and H with at most n vertices each, the isomorphism of $U_x(G)$ and $U_y(H)$ surely follows from the isomorphism of these rooted trees truncated at depth T . Motivated by applications in theory of distributed computing, Norris [Discrete Appl. Math. 1995] asks if $T(n) \leq n$. We answer this question in the negative by establishing that $T(n) = (2 - o(1))n$. Our solution uses basic tools of finite model theory such as a bisimulation version of the Immerman-Lander 2-pebble counting game.

The graphs G_n and H_n we construct to prove the lower bound for $T(n)$ also show some other tight lower bounds. Both having n vertices, G_n and H_n can be distinguished in 2-variable counting logic only with quantifier depth $(1 - o(1))n$. It follows that *color refinement*, the classical procedure used in isomorphism testing and other areas for computing the coarsest equitable partition of a graph, needs $(1 - o(1))n$ rounds to achieve color stabilization on each of G_n and H_n . Somewhat surprisingly, this number of rounds is not enough for color stabilization on the disjoint union of G_n and H_n , where $(2 - o(1))n$ rounds are needed.

1 Introduction

A homomorphism from a connected graph H onto a graph G is called a *covering map* if it is a bijection in the neighborhood of each vertex of H . In this case, we say that H is a *cover* of G or that H *covers* G . Given a vertex x of G , let $U_x(G)$ denote the unfolding of G into a tree starting from x . This tree is called the *universal cover* of G because $U_x(G)$ covers every cover H of G ; see examples in Fig. 1.

*Wilhelm-Schickard-Institut, Universität Tübingen, Sand 13, 72076 Tübingen, Germany.

†Humboldt-Universität zu Berlin, Institut für Informatik, Unter den Linden 6, D-10099 Berlin.
Supported by DFG grant VE 652/1–1. On leave from the Institute for Applied Problems of Mechanics and Mathematics, Lviv, Ukraine.

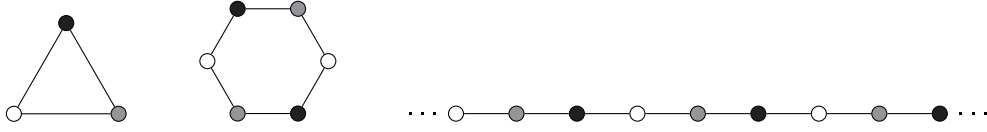


Figure 1: The cycle C_6 covers the cycle C_3 . The infinite path is the universal cover of both cycles. The corresponding covering maps are visualized by colors of vertices.

These notions appeared in algebraic and topological graph theory [6, 12, 29, 37], where they are helpful, for instance, in factorization of the characteristic polynomial of a graph [38] or in classification of projective planar graphs [31]. Further applications were found in such diverse areas as finite automata theory [30, 33], combinatorial group theory [39, 40], finite model theory [34], construction of expander graphs [1, 7] and, maybe most noticeably, in distributed computing [2, 8, 16, 42] (see also the surveys [26, 28]). The problem we consider in this paper arose in the last area.

In theory of distributed systems, a synchronous network of anonymous processors is presented as a graph G where, in a unit of time, two processors exchange messages if the corresponding vertices are connected by an edge. The processors are supposed to have unlimited computational power (for example, being automata with unbounded number of states). The processors at vertices of the same degree execute the same program. They are identical at the beginning, but later can have different states if they receive different messages from their neighbors. All processors have a common goal of arriving at a specified configuration of their states. For example, the *leader election* problem is to ensure that exactly one of the processors comes in the distinguished state “elected” while all others come in the “unelected” state.

For an integer $t \geq 0$, let $U_x^t(G)$ be the truncation of $U_x(G)$ at depth t . Two processors x and y are indistinguishable by their states up until time t if $U_x^t(G) \cong U_y^t(G)$, where \cong denotes isomorphism of rooted trees. In particular, none of x and y can be elected in time t as a leader. Moreover, leader election is possible only if G contains a vertex x such that $U_x(G) \not\cong U_y(H)$ for any other y (see, e.g., the discussion in [9]).

Another archetypical problem in distributed computing is *network topology recognition*, that is, identification of the isomorphism type of the underlying graph G or, at least, checking if G has a specified property \mathcal{P} . Angluin [2] observed that, if H covers G , then these two graphs are indistinguishable in the above model of distributed computation (and in other natural models of local computations). It follows that a class of graphs \mathcal{P} is recognizable only if it is closed under coverings. This basic observation is used, for example, in [11] where it is shown that, except for a few special cases, minor-closed classes of graphs are not closed under covers and, hence, cannot be recognized.

Two networks G and H cannot be distinguished in time t by the states of their processors $x \in V(G)$ and $y \in V(H)$ if $U_x^t(G) \cong U_y^t(H)$. Suppose that each of G and H has at most n nodes. What is the minimum time $T = T(n)$ such that $U_x^T(G) \cong U_y^T(H)$ surely implies $U_x(G) \cong U_y(H)$? In other words, $T(n)$ is the minimum time that suffices to distinguish x and y whenever at all possible.

Norris [33] considers the case when $G = H$ and shows that then $T(n) \leq n - 1$, improving upon an earlier bound of n^2 in [42] and thereby increasing the performance of several distributed algorithms. In the same paper, she asks if $T(n) \leq n$ in the general case of two graphs G and H . We answer this question in the negative by establishing that $T(n) = (2 - o(1))n$. The upper bound $T(n) < 2n$ is standard, and our main contribution is a construction of graphs G and H showing a lower bound $T(n) \geq 2n - 16\sqrt{n}$.

The bound of $2n$ is a standard upper bound for the communication round complexity of the distributed algorithms that are based on computing the isomorphism type of the universal cover $U_x(G)$ or on related concepts; see Tani [41]. Our result implies that this bound is tight up to a term of $o(n)$ for any algorithm that aims at gaining all knowledge about the network G available to a particular party x .

Our solution of Norris’s problem uses, perhaps for the first time in the area of distributed computing, a conceptual framework that was created in finite model theory. Specifically, our proof of the lower bound for $T(n)$ makes use of a bisimulation version of the 2-pebble counting game [24], that was used for diverse purposes in [19, 3]. The interplay between the two areas is discussed in more detail below.

Note that Norris in [33] considers directed graphs and allows multiple edges and loops. This setting is more general and as well important for modelling of distributed systems: Undirected graphs we consider in this paper just correspond to networks with bidirectional communication channels. It should be stressed that our lower bound for $T(n)$, while shown for undirected graphs, holds true in the setting of [33] by considering a simple orientation of the constructed graphs; see Remark 3.4.

Relation to color refinement. A coloring of the vertex set of an undirected graph G is called *stable* if any two equally colored vertices have the same number of neighbors of each color. In algebraic graph theory [18], the corresponding partition of $V(G)$ is called *equitable*. The well-known *color refinement* procedure begins with a uniform coloring of $V(G)$ and refines it step by step so that, if two vertices currently have equal colors but differently colored neighborhoods, then their new colors are different. The procedure terminates as soon as no further refinement is possible. The coloring obtained in this way is stable and gives a unique coarsest equitable partition of $V(G)$. More efficient implementations of color refinement have been developed in the literature starting from Hopcroft’s work on minimization of finite automata [23]; see [5] and references therein. Even in the simplest version, color refinement is used in isomorphism testing as a very efficient way to compute a canonical labeling for almost all input graphs (Babai, Erdős, and Selkow [4]). Angluin [2] uses color refinement to decide if two given graphs have a common cover. The last condition is important because, as follows from the discussion above, such graphs are indistinguishable by local computations.

Using the relationship between universal covers and color refinement observed in [2], our bound for $T(n)$ can be recasted as a result about the basic color refinement procedure. Let $Stab(G)$ denote the number of refinements made on the input G till stabilization. The graphs G_n and H_n we construct to solve Norris’s problem show that the obvious upper bound $Stab(G) < n$, where n is the number of vertices in G , is tight. Both G_n and H_n have n vertices, and both $Stab(G_n)$ and $Stab(H_n)$ are

$(1 - o(1))n$. Moreover, the disjoint union $G_n \cup H_n$ of these graphs demonstrates a counterintuitive phenomenon: $\text{Stab}(G_n \cup H_n) = (2 - o(1))n$, which means that stabilization of the disjoint union can require much more refinement rounds than stabilization of each component.

Relation to two-variable counting logic. As it is already mentioned, our main technical tool is a version of the 2-pebble counting game introduced by Immerman and Lander [24] for analysis of the expressibility of $\text{FO}_{\#}^2$, first-order logic with two variables and counting quantifiers. Let $D(G)$ denote the minimum quantifier depth of a formula defining the equivalence type of a graph G in $\text{FO}_{\#}^2$. The connection to color refinement established by Immerman and Lander in [24] can be used to show that $D(G) \leq \text{Stab}(G) + 2 \leq n + 1$ for G with n vertices. Our construction yields graphs asymptotically attaining this bound. While a linear lower bound of $D(G) \geq n/2 - O(1)$ can be easily seen by considering $G = P_n$, the path graph on n vertices, the exact asymptotics of the maximum value of $D(G)$ is determined for the first time. Note also that the parameter $D(G)$ is very small in the average case. As follows from [4], $D(G) \leq \text{Stab}(G) + 2 \leq 4$ for almost all G on n vertices.

Related work. A *fibration* is an analog of a covering for directed graphs, possibly with loops and multiple edges, where the local bijectivity property is required only on *in*-arrows. Boldi and Vigna [9] discuss the Norris’s question for fibrations and notice that the lower bound $T(n) \geq 2n - 2$ can in this setting be obtained by considering the following simple example of digraphs G and H along with vertices $x \in V(G)$ and $y \in V(H)$. Both G and H are obtained from the bidirectional path on the vertices $1, 2, \dots, n$ by cloning the arrow from $n - 1$ to n in G and by adding a loop at n in H ; furthermore, $x = 1$ and $y = 1$. The overall idea behind our construction is actually similar, though its implementation without loops and multiple edges (and with the bijectivity constraint on the entire neighborhood of a vertex) is not so simple.

In a more realistic model of a bidirectional network, an undirected graph G is endowed with a *port-numbering*. This is a labeling that determines an order on the set of the incident edges for each node. Hendrickx [22] proves in this setting¹ that $U_x^t(G) \cong U_y^t(G)$ implies $U_x(G) \cong U_y(G)$ for $t = O(d + d \log(n/d))$ where d denotes the diameter of G . This bound can be preferable to Norris’s bound of $n - 1$ if $d = o(n)$. The optimality of this bound is shown by Dereniowski, Kosowski, and Pajak [14]. Note that our solution of Norris’s problem can be extended also to port-numbered graphs; see Remark 3.5.

The results on the computational complexity of deciding if H covers G for two given graphs are surveyed in [15].

Organization of the paper. The connection between universal covers and color refinement is explored in Section 2. A key technical role is here played by Lemma 2.5, which is a kind of a reconstructibility result for rooted trees (cf. [25, 32]). Once this

¹In fact, in place of the universal covers Hendrickx [22] considers the related concept of a *view* introduced by Yamashita and Kameda in [42]. While the vertices of the universal cover $U_x(G)$ can be identified with the non-backtracking walks in G starting at x , the vertices of the view of $x \in V(G)$ correspond to all (not necessarily non-backtracking) walks in G starting at x . Isomorphism of truncated views is equivalent to isomorphism of truncated universal covers because both can be characterized by color refinement as in Lemma 2.6.

connection is established (Lemma 2.6), it readily yields the upper bound $T(n) < 2n$ (Lemma 2.7). The lower bound for $T(n)$ is obtained in Section 3 (Theorem 3.2). The proof uses the relationship between color refinement and the bisimulation version of the 2-pebble Immerman-Lander game [24]. The core of the proof is a construction of n -vertex graphs G and H containing vertices u and v respectively such that, while Spoiler can win starting from the position (u, v) , Duplicator can resist during $(2 - o(1))n$ rounds. In Section 4 we analyse the original version of the Immerman-Lander game on the same graphs and determine the maximum quantifier depth needed to define the equivalence type of a graph on n vertices in $\text{FO}_{\#}^2$ (Theorem 4.6). Theorem 4.5 in the same section is obtained as a by-product and stated for expository purposes. It shows another connection between universal covers and logic: Deciding if given graphs have a common cover is P-hard by a reduction from the $\text{FO}_{\#}^2$ -equivalence problem, whose P-hardness is established by Grohe [20]. This reduction is implicitly contained in [36, Theorem 2.2]; see also [3]. We conclude with some open questions in Section 5.

2 Universal covers and color refinement

2.1 Basic definitions and facts

Unless stated otherwise, we consider unlabeled undirected graphs without loops and multiple edges. Given a graph G , we denote its vertex set by $V(G)$. The *neighborhood* of a vertex v consists of all vertices adjacent to v and is denoted by $N(v)$.

In this section we consider, along with finite, also infinite graphs. The following definitions, which are standard in the finite case, apply as well to infinite graphs. A graph is *connected* if any two its vertices can be connected by a (finite) path. A graph is *acyclic* if it contains no (finite) cycle. A (possibly infinite) *tree* is an acyclic connected graph.

All graphs in this section are supposed to be **connected**. Let α be a homomorphism from H onto G . If α is a bijection from $N(v)$ onto $N(\alpha(v))$ for each $v \in V(H)$, then it is called a *covering map*, and H is called a *covering graph* (or a *cover*) of G . Sometimes we also say that H *covers* G . Note that the following fact holds true both for finite and infinite graphs.

Lemma 2.1. *If α is a covering map from a connected graph H onto a tree T , then H is a tree and α is an isomorphism from H to T .*

Proof. If H had a cycle, α would take it to a closed non-backtracking walk in T , which is impossible because such a walk must contain a cycle. Next we show that α is injective. Fix a vertex x in H and let α^i denote the restriction of α to the vertices of H at the distance at most i from x . For every $i \geq 0$, α^i is injective; this follows by induction on i . Therefore, α is injective. Finally, note that an injective covering map is an isomorphism because it always takes a pair of non-adjacent vertices to a pair of non-adjacent vertices. ■

Given a connected graph G and a vertex $x \in V(G)$, define a graph $U_x(G)$ as follows. The vertex set of $U_x(G)$ consists of non-backtracking walks in G starting at x , that is, of sequences (x_0, x_1, \dots, x_k) such that $x_0 = x$, x_i and x_{i+1} are adjacent, and $x_{i+1} \neq x_{i-1}$. Two such walks are adjacent in $U_x(G)$ if one of them extends the other by one component, that is, one is $(x_0, \dots, x_k, x_{k+1})$ and the other is (x_0, \dots, x_k) . Notice the following properties of this construction.

Lemma 2.2.

1. $U_x(G)$ is a tree.
2. The map γ_G defined by $\gamma_G(x_0, \dots, x_k) = x_k$ is a covering map from $U_x(G)$ to G .
3. If α is a covering map from H onto G , then there is a covering map β from $U_x(G)$ onto H such that $\gamma_G = \alpha \circ \beta$.

Proof. We skip Parts 1 and 2 that can be shown by a direct argument. To prove Part 3, fix a vertex y of H such that $\alpha(y) = x$. Define $\bar{\alpha}(y_0, \dots, y_k) = (\alpha(y_0), \dots, \alpha(y_k))$ and note that $\bar{\alpha}$ is a covering map from $U_y(H)$ to $U_x(G)$. Since the two graphs are trees, $\bar{\alpha}$ is an isomorphism by Lemma 2.1. Note that $\gamma_G \circ \bar{\alpha} = \alpha \circ \gamma_H$. Thus, we can set $\beta = \gamma_H \circ \bar{\alpha}^{-1}$. ■

Call U a *universal cover* of G if U covers any covering graph of G . Lemma 2.2.3 implies that $U_x(G)$ is a universal cover of G . The next lemma shows that we could define the universal cover of G , uniquely up to isomorphism, as a tree covering G .

Lemma 2.3.

1. All universal covers of G are isomorphic trees.
2. Any tree covering G is a universal cover of G .

Proof. 1. By definition, all universal covers of G cover each other. Since one of them, namely $U_x(G)$, is a tree, the claim follows from Lemma 2.1.

2. Assume that a tree T covers G . Let U be a universal cover of G . Then U covers T , and the two graphs are isomorphic by Lemma 2.1. Like U , the tree T is therefore a universal cover of G . ■

Lemma 2.4.

1. If C covers G , then their universal covers U_C and U_G are isomorphic.
2. If G and H have a common covering graph, then their universal covers U_G and U_H are isomorphic.

Proof. 1. By Lemma 2.3.1, U_C is a tree. Since U_C covers C and C covers G , U_C covers G . By Lemma 2.3.2, U_C is a universal cover of G and, therefore, is isomorphic to U_G .

2. Consider a common cover C of G and H and apply Part 1. ■

Lemma 2.4 shows that two connected graphs have a common covering graph if and only if their universal covers are isomorphic. It is known (Leighton [27]) that graphs have a common covering graph if and only if they have a common *finite* covering graph.

2.2 A reconstruction lemma for rooted trees

A *rooted tree* T_v is a tree with one distinguished vertex v , which is called *root*. An isomorphism of rooted trees should not only preserve the adjacency relation but also map the root to the root. The *depth* of T_v is the maximum distance from v to a leaf. We write T_v^i to denote the subgraph of T_v induced on the vertices at the distance at most i from v . Note that T_v^i inherits the root v .

If $v \in V(G)$, let $G - v$ denote the graph obtained by removing v from G along with all incident edges. If $w \in N(v)$, then $T_v(w)$ denotes the component of $T_v - v$ containing w . The tree $T_v(w)$ will be supposed to be rooted at w and will be referred to as a *branch* of T_v .

There are well-known results about the reconstructibility of the isomorphism type of a tree from the isomorphism types of some of its proper subgraphs [25, 32]. We will need a result of this kind showing that a rooted tree is reconstructible, up to isomorphism, from a family of related rooted trees of smaller depth.

Lemma 2.5. *Let T and S be trees and $x \in V(T)$ and $y \in V(S)$ be their vertices of the same degree with neighborhoods $N(x) = \{x_1, \dots, x_k\}$ and $N(y) = \{y_1, \dots, y_k\}$. Let $r \geq 1$. Suppose that $T_x^{r-1} \cong S_y^{r-1}$ and $T_{x_i}^r \cong S_{y_i}^r$ for all $i \leq k$. Then $T_x^{r+1} \cong S_y^{r+1}$.*

Proof. Note that, for each $i \leq k$, x_i and y_i have equal degrees. Let A_0, A_1, \dots, A_m be the branches of $T_{x_i}^r$ (that is, the components of $T_{x_i}^r - x_i$ rooted at the neighbors of x_i). The dependence on i in the notation A_j is dropped. Similarly, let B_0, B_1, \dots, B_m be the branches of $S_{y_i}^r$. We suppose that $A_0 = T_{x_i}^r(x)$ and $B_0 = S_{y_i}^r(y)$, that is, the roots of A_0 and B_0 are x and y respectively.

Our task is to prove that $A_0 \cong B_0$ for each $i \leq k$. Along with the conditions $T_{x_i}^r \cong S_{y_i}^r$, this will imply that $T_x^{r+1}(x_i) \cong S_y^{r+1}(y_i)$ for all $i \leq k$. Merging these isomorphisms with the map taking x to y and each x_i to y_i , we will obtain a desired isomorphism from T_x^{r+1} to S_y^{r+1} .

Consider an isomorphism α_i from $T_{x_i}^r$ to $S_{y_i}^r$. If α_i maps A_0 onto B_0 , we are done. Otherwise, without loss of generality we can assume that α_i maps A_0 onto B_1 , A_1 onto B_0 , and A_j onto B_j for every $2 \leq j \leq m$.

Since both A_0 and B_0 have depth at most $r - 1$, we need to show that $A_0^{r-1} \cong B_0^{r-1}$, where $A_0^h = (A_0)_x^h$ and similar notation is used also for the rooted tree B_0 . Using induction, we will prove that $A_0^h \cong B_0^h$ for all $h = 0, 1, \dots, r - 1$.

The base case of $h = 0$ is trivial. Note that it proves the lemma for $r = 1$. Let $r \geq 2$ and $h \leq r - 2$. Assume that $A_0^h \cong B_0^h$ and deduce from here that $A_0^{h+2} \cong B_0^{h+2}$ (for $h = r - 2$ note that $A_0^r = A_0^{r-1}$ and $B_0^r = B_0^{r-1}$). Via α_i we get $A_1^h \cong B_1^h$. Since we also have $A_j^h \cong B_j^h$ for all $j \geq 2$, we can combine these isomorphisms and obtain an isomorphism from $T_x^{h+2}(x_i)$ onto $S_y^{h+2}(y_i)$; see Fig. 2.

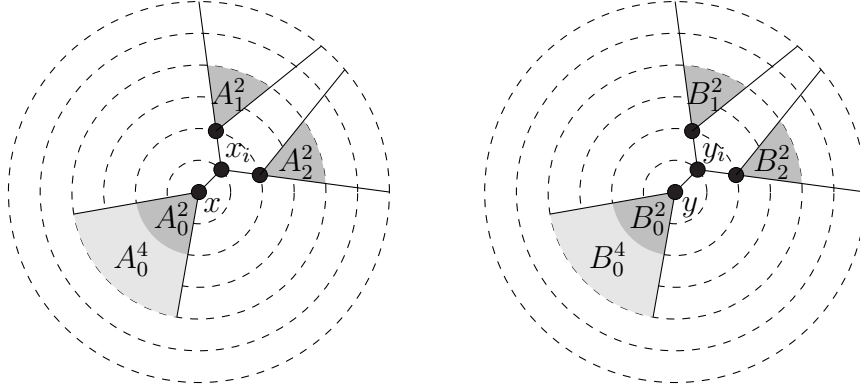


Figure 2: Proof of Lemma 2.5: The inductive step from $h = 2$ to $h + 2 = 4$ ($A_0^2 \cong B_0^2 \implies A_0^4 \cong B_0^4$). The dashed circles correspond to the metric in T_x and S_y . Note that $T_x(x_i)$ is rooted at x_i and has branches A_1^2 and A_2^2 .

The condition $T_x^{r-1} \cong S_y^{r-1}$ implies that $T_x^{h+2} \cong S_y^{h+2}$. Since T_x^{h+2} actually consists of $T_x^{h+2}(x_i)$ and A_0^{h+2} and S_y^{h+2} consists of $S_y^{h+2}(y_i)$ and B_0^{h+2} , we conclude that A_0^{h+2} and B_0^{h+2} are also isomorphic. \blacksquare

2.3 A relationship between universal covers and color refinement

A partition Π of the vertex set of a graph G is called *equitable* if for any elements $X \subseteq V(G)$ and $Y \subseteq V(G)$ of Π the following is true: Any two vertices u and v in X have the same number of neighbors in Y . A trivial example of an equitable partition is the partition of $V(G)$ into singletons. There is a unique equitable partition Π^* that is the *coarsest* in the sense that any other equitable partition Π is a subpartition of Π^* . This partition can be found by the following *color refinement* procedure. Define a sequence of colorings C^i of G recursively. The initial coloring C^0 is uniform. Then,

$$C^{i+1}(u) = \langle C^i(u), \{\{ C^i(a) : a \in N(u) \}\} \rangle, \quad (1)$$

where $\langle \dots \rangle$ denotes an ordered pair and $\{\{ \dots \}\}$ denotes a multiset. Regard C^i as a partition of $V(G)$ (consisting of the monochromatic classes of vertices) and note that C^{i+1} is a refinement of C^i . It follows that, eventually, $C^{s+1} = C^s$ for some s ; hence, $C^i = C^s$ for any $i \geq s$. Such a partition C^s is called *stable*. It is easy to see that C^s is equitable, and an inductive argument shows that it is the coarsest; see [10, Lemma 1]. The minimum s such that C^s is stable will be denoted by $\text{Stab}(G)$.

We now show that the color $C^i(v)$ describes the isomorphism type of the universal cover $U_v(G)$ truncated at depth i .

Lemma 2.6. *Let U and W be universal covers of graphs G and H respectively. Furthermore, let α be a covering map from U to G and β be a covering map from W to H . Then $U_x^i \cong W_y^i$ if and only if $C^i(\alpha(x)) = C^i(\beta(y))$.*

Proof. We use induction on i . The base case of $i = 0$ is trivial.

Assume that the claim is true for i . By Lemma 2.5, $U_x^{i+1} \cong W_y^{i+1}$ if and only if $U_x^i \cong W_y^i$ and $\{\{U_z^i : z \in N(x)\}\} = \{\{W_z^i : z \in N(y)\}\}$, where the multisets consist of the isomorphism types of rooted trees. By the induction assumption, the former condition is equivalent to $C^i(\alpha(x)) = C^i(\beta(y))$ and the latter condition is equivalent to $\{\{C^i(\alpha(z)) : z \in N(x)\}\} = \{\{C^i(\beta(z)) : z \in N(y)\}\}$. Since α and β are bijective on $N(x)$ and $N(y)$, the last equality can be rewritten as $\{\{C^i(a) : a \in N(\alpha(x))\}\} = \{\{C^i(a) : a \in N(\beta(y))\}\}$. By the definition of $C^{i+1}(u)$, we conclude that $U_x^{i+1} \cong W_y^{i+1}$ if and only if $C^{i+1}(\alpha(x)) = C^{i+1}(\beta(y))$. ■

The next lemma shows an upper bound for the truncation depth sufficient to detect isomorphism of universal covers.

Lemma 2.7. *Suppose that G and H are connected graphs with at most n vertices each. Let U and W be their universal covers. Then $U_x \cong W_y$ if and only if $U_x^{2n-1} \cong W_y^{2n-1}$.*

Proof. One direction is trivial. Suppose that $U_x^{2n-1} \cong W_y^{2n-1}$. Fix covering maps α of U onto G and β of W onto H . By Lemma 2.6, $C^{2n-1}(\alpha(x)) = C^{2n-1}(\beta(y))$. Note that C^{2n-1} is a stable partition of the disjoint union of G and H . It follows that $C^i(\alpha(x)) = C^i(\beta(y))$ for all i . Using Lemma 2.6 again, we conclude that $U_x^i \cong W_y^i$ for all i . Consider the set of all isomorphisms from U_x^i to W_y^i , $i \geq 0$. Define an auxiliary graph on this set joining an isomorphism ϕ from U_x^i to W_y^i and an isomorphism ψ from U_x^{i+1} to W_y^{i+1} by an edge if ψ extends ϕ . By König's lemma, this graph contains an infinite path $\phi_0 \subset \phi_1 \subset \phi_2 \subset \dots$. The union of these isomorphisms gives us an isomorphism from U_x to W_y . ■

2.4 Existence of a common cover: Angluin's algorithm

Lemma 2.8. *Suppose that connected graphs G and H have at most n vertices each. Then the following conditions are equivalent:*

- (1) G and H have a common covering graph.
- (2) $\{C^{2n-1}(u) : u \in V(G)\} \cap \{C^{2n-1}(v) : v \in V(H)\} \neq \emptyset$.
- (3) $\{C^{2n-1}(u) : u \in V(G)\} = \{C^{2n-1}(v) : v \in V(H)\}$.
- (4) $\{C^{s+1}(u) : u \in V(G)\} = \{C^{s+1}(v) : v \in V(H)\}$, where $s = \text{Stab}(G)$.

Note that Condition 4 improves Condition 3 because $s + 1 \leq n$. While the equivalence to Conditions 2 and 3 means that Condition 1 can be detected after stabilization of the coloring C^i on the disjoint union of G and H , the equivalence to Condition 4 means that this can actually be done as soon as C^i stabilizes at least on *one* of the graphs G and H . Another consequence of Lemma 2.8 is this: If G and H do not share a covering graph, then the color sets $\{C^i(u) : u \in V(G)\}$ and $\{C^i(v) : v \in V(H)\}$ must be unequal starting from $i = n$ and disjoint starting from $i = 2n - 1$.

Proof. (1) \iff (2). By Lemma 2.4, G and H have a common cover if and only if their universal covers U and W are isomorphic. Note that $U \cong W$ iff $U_x \cong W_y$ for some x and y . By Lemmas 2.7 and 2.6,

$$U_x \cong W_y \iff U_x^{2n-1} \cong W_y^{2n-1} \iff C^{2n-1}(\alpha(x)) = C^{2n-1}(\beta(y)),$$

for covering maps α from U to G and β from W to H . It follows that G and H have a common cover iff $C^{2n-1}(u) = C^{2n-1}(v)$ for some $u \in V(G)$ and $v \in V(H)$, which is exactly Condition 2 in the lemma.

(2) \iff (3). Condition 3 implies Condition 2 and follows by Lemma 2.6 from the isomorphism $U \cong W$, which is implied by Condition 2.

(3) \iff (4). Consider C^s , which is stable on G . If the sets $\{C^s(u) : u \in V(G)\}$ and $\{C^s(v) : v \in V(H)\}$ are not equal, then Conditions 3 and 4 are both false because unequal vertex colors cannot become equal later. Suppose that these sets are equal and compare the partitions C^s and C^{s+1} of $V(G) \cup V(H)$. If $C^s \neq C^{s+1}$, this means that the number of colors in the C^{s+1} -coloring of H is strictly larger than in the C^{s+1} -coloring of G . Therefore, Condition 4 is false, and Condition 3 is false too. If $C^s = C^{s+1}$, this means that the partition stabilizes on $V(G) \cup V(H)$. In this case, vertices equally colored with respect to C^s remain equally colored with respect to any C^i , which implies that both Conditions 3 and 4 are true. ■

Lemma 2.8 justifies Angluin's algorithm [2] for deciding in polynomial time whether or not two given graphs have a common covering graph. Any of Conditions 2, 3, or 4 can be checked efficiently. Note that we cannot compute the colors $C^{2n-1}(u)$ literally as they defined by Equation (1) because the length of $C^i(u)$ grows exponentially with i . We can overcome this complication by renaming the colors after each refinement step (doing so, we will never need more than $2n$ color names).

3 Norris's problem

In order to characterize expressibility of two-variable logic with counting quantifiers, Immerman and Lander [24] introduced the *2-pebble counting game* on graphs G and H . We need a restricted version of this game that was defined and used for diverse purposes in [19, 3]. We call it the *counting bisimulation game*. The game is played by two players, Spoiler and Duplicator, to whom we will refer as he and she respectively. Each player has a pair of distinct pebbles p and q , that are put on vertices of G and H during the game. The two copies of the same pebble are always in different graphs. In each *round* of the game the copies of one of the pebbles p and q change their positions while the copies of the other pebble do not move. A round where the pebble p moves is played as follows. First, Spoiler removes the copies of p from the board (if they are already put there). Suppose that the copies of q are located on vertices $u \in V(G)$ and $v \in V(H)$ at this point of time. Then Spoiler specifies a set of vertices A in one of the graphs such that either $A \subseteq N(u)$ or $A \subseteq N(v)$. Duplicator has to respond with a set of vertices B in the other graph such that $B \subseteq N(v)$ or $B \subseteq N(u)$ respectively. Duplicator must keep the condition $|B| = |A|$

true; otherwise she loses the game in this round. Finally, Spoiler puts p on a vertex $b \in B$. In response Duplicator has to place the other copy of this pebble on a vertex $a \in A$. In the next round the roles of p or q are interchanged. We write $\text{GAME}^r(G, u, H, v)$ to denote the r -round counting bisimulation game on G and H starting from the position where the vertices $u \in V(G)$ and $v \in V(H)$ are occupied by one of the pebbles, say, by the copies of q . Duplicator wins this game if she does not lose any of the r rounds.

Lemma 3.1. *Duplicator has a winning strategy in $\text{GAME}^i(G, u, H, v)$ if and only if $C^i(u) = C^i(v)$.*

Proof. We use induction on i . For $i = 0$ the claim is trivially true. Assume that it is true for some $i \geq 0$.

Suppose that $C^{i+1}(u) = C^{i+1}(v)$. Then Duplicator wins $\text{GAME}^{i+1}(G, u, H, v)$ as follows. Without loss of generality, suppose that in the first round Spoiler specifies a set $A \subseteq N(u)$. It follows from $C^{i+1}(u) = C^{i+1}(v)$ that

$$\{\{C^i(a) : a \in N(u)\}\} = \{\{C^i(b) : b \in N(v)\}\}. \quad (2)$$

This means that Duplicator can “mirror” the set A with a set $B \subseteq N(v)$ such that (2) stays true if $N(u)$ is restricted to A and $N(v)$ is restricted to B . This allows Duplicator to ensure pebbling vertices $a \in V(G)$ and $b \in V(H)$ such that $C^i(a) = C^i(b)$ and wins the remaining i rounds by the induction assumption.

Suppose now that $C^{i+1}(u) \neq C^{i+1}(v)$. Then Spoiler wins $\text{GAME}^{i+1}(G, u, H, v)$ as follows. The inequality of colors $C^{i+1}(u)$ and $C^{i+1}(v)$ means that either $C^i(u) \neq C^i(v)$ or $\{\{C^i(a) : a \in N(u)\}\} \neq \{\{C^i(b) : b \in N(v)\}\}$. In the former case Spoiler wins in i moves by the induction assumption. In the latter case, there is a C^i -color c such that for one of the two vertices, say, for u , the set of neighbors a with $C^i(a) = c$ is strictly larger than the analogous set for v . Specifying the larger set as A , Spoiler ensures pebbling a and b such that $C^i(a) \neq C^i(b)$ and wins in the next i moves by the induction assumption. ■

Theorem 3.2. *For each n , there are n -vertex graphs G and H whose universal covers U and W contain vertices x and y such that $U_x^i \cong W_y^i$ for all $i \leq 2n - 16\sqrt{n}$ while $U_x \not\cong W_y$.*

To prove this result, consider graphs $G_{s,t}$ and $H_{s,t}$ as shown in Fig. 3. Each of the graphs is a chain of t blocks, one head block and $t - 1$ tail blocks. All tail blocks in both $G_{s,t}$ and $H_{s,t}$ are copies of the same graph B_s with $s + 10$ vertices (where s is the number of vertices of degree at most 2 spanning a path in B_s). The head blocks of $G_{s,t}$ and $H_{s,t}$ are different. Note that the head block of $G_{s,t}$ also contains a copy of B_s . Both head blocks have $2s + 15$ vertices. Thus, both $G_{s,t}$ and $H_{s,t}$ have

$$n = (t + 1)(s + 10) - 5$$

vertices. Both graphs have a single vertex of degree 1; we denote these vertices by u and v respectively.

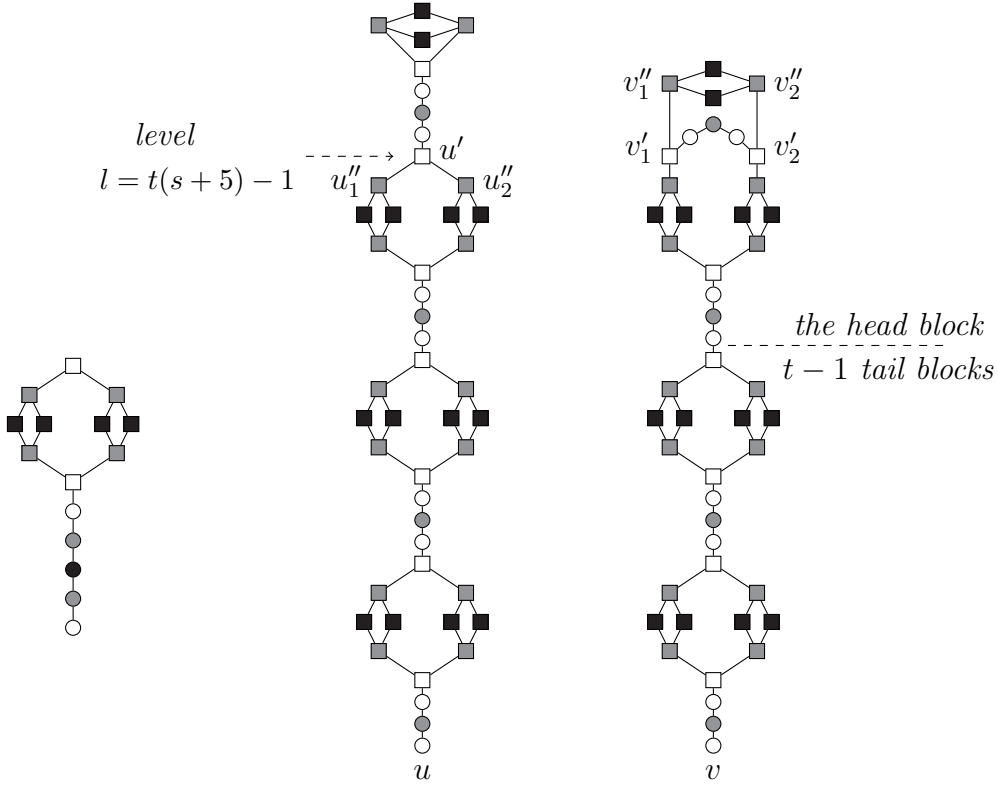
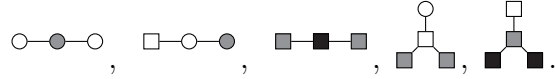


Figure 3: The tail block B_s for $s = 5$ and the graphs $G_{s,t}$ and $H_{s,t}$ for $s = 3$, $t = 3$.

The graphs $G_{s,t}$ and $H_{s,t}$ are uncolored. However, we distinguish $\lceil s/2 \rceil + 3$ types of vertices in them, that are presented in Fig. 3 by auxiliary colors and shapes.

We will use the following properties of $G_{s,t}$ and $H_{s,t}$.

- (A) The partition of $V(G_{s,t}) \cup V(H_{s,t})$ by types is almost equitable: With the exception of u and v , any two vertices of the same type have the same number of neighbors of each type. For example, if $s = 3$, all possible neighborhoods are these:



- (B) Given a vertex z in $G_{s,t}$ or $H_{s,t}$, we define its *level* $\ell(z)$ as the distance from z to u or v respectively. Up to the level

$$l = t(s + 5) - 1, \quad (3)$$

the following holds true: All vertices at the same level have the same type.

Lemma 3.3. *Consider the counting bisimulation $\text{GAME}^r(G_{s,t}, u, H_{s,t}, v)$.*

1. *Spoiler has a winning strategy for $r = 2l + 1$;*
2. *Duplicator has a winning strategy for any $r \leq 2l$, where l is defined by (3).*

Proof. 1. Let $a_0 = u$ and $b_0 = v$. As long as $i \leq l$, in the i -th round Spoiler specifies A consisting of a single vertex a_i such that a_i and a_{i-1} are adjacent and $\ell(a_i) = \ell(a_{i-1}) + 1$. Let $B = \{b_i\}$ be Duplicator's response. As long as $\ell(b_i) = \ell(b_{i-1}) + 1$, we have

$$\ell(a_i) = \ell(b_i) = i. \quad (4)$$

This equality can be broken only if $\ell(b_i) = \ell(b_{i-1}) - 1$. In this case $\ell(b_i) = i - 2$ while $\ell(a_i) = i$, and Spoiler wins in the other $i - 1$ moves by pebbling vertices in $H_{s,t}$ along a shortest path from b_i to v . He needs $i - 2$ moves to reach v and wins in one extra move because the corresponding vertex in $G_{s,t}$ has degree larger than 1.

We, therefore, assume that the condition (4) holds true for all $i \leq l$. In the notation of Fig. 3, this means that $a_l = u'$ and $b_l = v'_j$ for $j = 1$ or $j = 2$. In the $(l + 1)$ -th round Spoiler specifies the set $A = \{u''_1, u''_2\}$ consisting of the two \blacksquare -neighbors of u' (both at the level $l - 1$). If Duplicator's response B contains the \circ -neighbor of v'_j , Spoiler pebbles it and wins in the next round because the vertices pebbled in $G_{s,t}$ and $H_{s,t}$ have different degrees. Otherwise, B consists of the two \blacksquare -neighbors of v'_j . One of them, v''_j , is at the level $l + 1$. Spoiler pebbles it, while Duplicator has to pebble u''_k for $k = 1$ or $k = 2$. Using the fact that $\ell(u''_k) = l - 1$ and $\ell(v''_j) = l + 1$, Spoiler wins in the next l rounds by pebbling vertices along a shortest path from u''_k to u . At total, he makes $2l + 1$ moves to win.

2. As before, let a_i and b_i denote the vertices of $G_{s,t}$ and $H_{s,t}$ pebbled in the i -th round. Note that Spoiler can win in the next $(i + 1)$ -th round only if a_i and b_i have different degrees. A sufficient condition for a_i and b_i having equal degrees is that these vertices are of the same type (excepting the case that one of them is u or v). Property A of $G_{s,t}$ and $H_{s,t}$ implies that, if a_{i-1} and b_{i-1} have the same type, Duplicator can ensure the same for a_i and b_i unless $a_{i-1} = u$ or $b_{i-1} = v$ (the case that both $a_{i-1} = u$ and $b_{i-1} = v$ is also favorable for Duplicator). This observation is the basis of Duplicator's strategy. In the first phase of the game, Duplicator keeps the levels of a_i and b_i equal. Up to level l , by property B this implies also the equality of types.

More precisely, we fix a strategy for Duplicator such that

1. $\ell(a_i) = \ell(b_i)$ as long as

$$\ell(a_{i-1}) = \ell(b_{i-1}) < l; \quad (5)$$

2. a_i and b_i have the same type as long as

$$\min\{\ell(a_{i-1}), \ell(b_{i-1})\} > 0 \text{ or } \ell(a_{i-1}) = \ell(b_{i-1}) = 0. \quad (6)$$

Duplicator can keep Conditions 1 and 2 true due to Properties A and B of $G_{s,t}$ and $H_{s,t}$. Note that (5) implies (6) and that fulfilling Condition 1 implies also fulfilling Condition 2. Thus, Duplicator is alive as long as (6) holds true.

Note now that (5) can be broken for $i - 1 = l$ at the earliest (i.e., not earlier than in the l -th round) and, when this happens, we will have $\ell(a_{i-1}) = \ell(b_{i-1}) = l$. Starting from this point, Spoiler needs no less than l moves to break (6). Therefore, Duplicator survives at least during the first $2l$ rounds irrespectively of Spoiler's strategy. \blacksquare

Turning back to the proof of Theorem 3.2, consider $G = G_{s,t}$ and $H = H_{s,t}$, where we set $s = 2t + 1$. Thus, both G and H have $n = (t + 1)(s + 10) - 5 = 2t^2 + 13t + 6$ vertices. By Lemma 3.3.2, Duplicator has a winning strategy in the counting bisimulation $\text{GAME}^r(G, u, H, v)$ for

$$r = 2t(s + 5) - 2 = 4t^2 + 12t - 2 = 2n - 14t - 14 > 2n - 16\sqrt{n}.$$

By Lemma 3.1, $C^r(u) = C^r(v)$. Consider covering maps α from U to G and β from W to H . Take x and y such that $\alpha(x) = u$ and $\beta(y) = v$. By Lemma 2.6, $U_x^r \cong W_y^r$. On the other hand, combining Lemmas 2.6 and 3.1 with Lemma 3.3.1, we conclude that $U_x^{r+1} \not\cong W_y^{r+1}$, which implies that $U_x \not\cong W_y$. This proves the theorem in the case that $n = 2t^2 + 13t + 6$ for some t .

For any other n , fix t such that $2t^2 + 13t + 6 < n < 2(t+1)^2 + 13(t+1) + 6$. Now we consider G obtained from $G_{2t+1,t}$ by adding new $k = n - (2t^2 + 13t + 6)$ vertices and connecting each of them to u by an edge. The graph H is obtained from $H_{2t+1,t}$ in the same way. The new vertices do not affect the outcome of $\text{GAME}^r(G, u, H, v)$ if $k > 3$. If $k \leq 3$, Duplicator can resist at most one round longer. Since $r + 1 = 4t^2 + 12t - 1$ is still larger than $2n - 16\sqrt{n}$, the proof is complete.

Remark 3.4. Theorem 3.2 is true also for oriented graphs. To see this, let us orient the graphs G and H constructed in the proof. Note that adjacent vertices in these graphs have different types. Fix arbitrarily an order on the set of all vertex types. For each pair of adjacent vertices a and b , we draw an arrow from a to b if the type of b has higher position with respect to this order than the type of a , for example,

$$\bullet \rightarrow \circ, \quad \circ \rightarrow \square, \quad \square \rightarrow \blacksquare, \quad \blacksquare \rightarrow \blacksquare. \quad (7)$$

(assuming $s = 3$ as in Fig. 3). Denote the corresponding oriented graphs by G' and H' . Let $U = U_u(G)$ and $W = U_v(H)$. The vertices of U and W inherit the types in a natural way. Orient U and W accordingly to (7) and denote the resulting oriented graphs by U' and W' . Note that U' and W' are universal covers of G' and H' respectively. Furthermore, isomorphism of U and W truncated at depth i implies isomorphism of U' and W' truncated at the same depth.

Remark 3.5. With an additional effort, an analog of Theorem 3.2 can be obtained also for port-numbered graphs, which are a popular model of distributed networks [2, 14, 22, 42]. Formally, a port-numbered graph can be defined as a relational structure that is an undirected graph G of maximum degree D where, for each edge $\{x, y\}$, each of the ordered pairs (x, y) and (y, x) satisfies exactly one of the binary relations B_1, \dots, B_D . For every vertex x of degree d , each of the relations B_1, \dots, B_d must be satisfied on some (x, y) . Thus, all arrows (x, y) emanating from x satisfy pairwise different relations B_i for $i \leq d$. The meaning of $B_i(x, y)$ is that the edge $\{x, y\}$ receives the number i among the edges incident to x . Since the techniques used in the proof of Theorem 3.2 easily generalize to arbitrary binary structures, we can proceed basically in the same way.

We first construct undirected graphs and then endow them with a suitable port-numbering. Specifically, we construct graphs $G_{s,t}$ and $H_{s,t}$ similarly to the proof

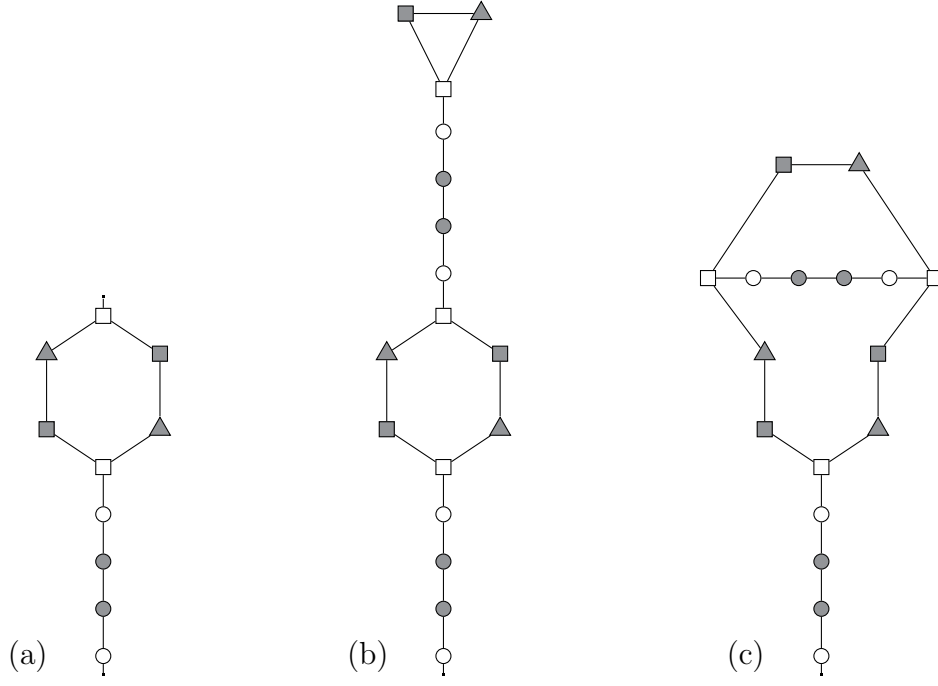


Figure 4: (a) The tail block B_4 . (b) The head block of $G_{4,t}$. (c) The head block of $H_{4,t}$.

of Theorem 3.2 from the tail and the head blocks depicted in Fig. 4. The types of vertices shown in the picture are now considered to be colors of vertices. Thus, $G_{s,t}$ and $H_{s,t}$ are vertex-colored graphs. We define $G = G_{2s,t}$ and $H = H_{2s,t}$; note that the first parameter is now even.

The modified construction ensures some special properties of the color partition that will be used below. The first useful property is the same as in the proof of Theorem 3.2: The partition is almost equitable (except for the bottom vertices u and v). This allows us to prove that

$$U_u^r(G) \cong U_v^r(H) \text{ for } r = 2n - O(\sqrt{n}). \quad (8)$$

On the other hand, $U_u(G) \not\cong U_v(H)$. In fact, even a stronger statement is true. Let G_0 and H_0 denote the uncolored versions of G and H respectively. Like in the proof of Theorem 3.2, we can prove that

$$U_u(G_0) \not\cong U_v(H_0). \quad (9)$$

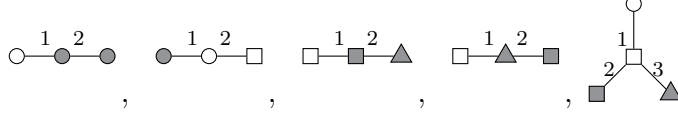
This actually follows from the fact that the eccentricity of the vertex u in G_0 is greater than the eccentricity of v in H_0 .

Now, we convert G and H into port-numbered graphs as follows. Fix an arbitrary order on the vertex colors, for example,

$$\circ, \bullet, \square, \blacksquare, \blacktriangle \quad (10)$$

(we assume $s = 4$ as in Fig. 4). Notice another useful property of the vertex coloring: The neighbors of each vertex x have pairwise different colors. This allows

us to unambiguously enumerate the edges incident to x consistently with the order (10). In our example, this results in the following labeling:



(furthermore, the single edge incident to u or to v is labeled by 1). As a result, each edge $\{x, y\}$ gets port labels that depend solely on the colors of x and y :

$$\begin{array}{c} \text{white} \text{---} \text{white} \\ \text{white} \text{---} \text{grey} \\ \text{grey} \text{---} \text{white} \\ \text{grey} \text{---} \text{grey} \\ \text{white} \text{---} \text{grey} \\ \text{grey} \text{---} \text{white} \end{array} \quad (11)$$

Denote the obtained port-numbered vertex-colored graphs by G' and H' and their uncolored versions by G'_0 and H'_0 .

Endow the universal covers $U = U_u(G)$ and $W = U_v(H)$ with port-numbering according to (11). Denote the resulting port-numbered trees by U' and W' and their uncolored versions by U'_0 and W'_0 . Note that U' and W' are the universal covers of G' and H' , and U'_0 and W'_0 are the universal covers of G'_0 and H'_0 respectively. Any isomorphism between U and W is an isomorphism between U' and W' and, hence, also between U'_0 and W'_0 . The same holds true as well for the truncations at any depth. By (8), this implies that $U_u^r(G'_0) \cong U_v^r(H'_0)$ for $r = 2n - O(\sqrt{n})$. Finally, $U_u(G'_0) \not\cong U_v(H'_0)$. This readily follows from (9).

Using Lemma 2.6, from Theorem 3.2 we derive the following fact.

Corollary 3.6. *For each n , there are n -vertex graphs G and H with disjoint vertex sets such that $\text{Stab}(G \cup H) = (2 - o(1))n$.*

4 Quantifier depth in two-variable counting logic

We consider first-order logic FO for graphs with two binary relations for adjacency and equality of vertices. Let FO^2 denote the fragment of FO consisting of formulas built from only two variables. We will enrich the language by using expressions of the type $\exists^m x \Psi(x)$ in order to say that there are at least m vertices x with property $\Psi(x)$. The *counting quantifier* \exists^m contributes 1 in the quantifier depth irrespectively of the value of m . The corresponding syntactic extension of FO^2 will be denoted by $\text{FO}_{\#}^2$. The following fact follows directly from the definition (1) by induction on i .

Lemma 4.1. *For any possible C^i -color c there is a formula $\Phi(x)$ in $\text{FO}_{\#}^2$ of quantifier depth i such that, for every graph G and its vertex u , $C^i(u) = c$ if and only if $\Phi(x)$ is true on G for $x = u$. ■*

We say that a formula Φ *distinguishes* G and H if it is true on exactly one of the graphs. Let $D(G, H)$ denote the minimum quantifier depth of such a formula in $\text{FO}_{\#}^2$.

The 2-pebble counting game [24] on graphs G and H differs from its bisimulation version described in the preceding section in that the sets A and B in each round are constrained only by the equality $|B| = |A|$. Instead, Duplicator must now ensure that the pebbling after each round determines a partial isomorphism, that is, the two pebbled vertices in G are adjacent/equal exactly if the pebbled vertices in H are adjacent/equal (in the bisimulation version, they were automatically not equal and adjacent in both graphs).

Lemma 4.2 (Immerman and Lander [24]). *$D(G, H) \leq r$ if and only if Spoiler has a winning strategy in the r -round counting game on G and H . ■*

Lemma 4.3. *Suppose that at least one of graphs G and H is connected. Let $s = \text{Stab}(G)$. Then G and H are indistinguishable in $\text{FO}_{\#}^2$ if and only if these graphs have equal number of vertices and*

$$\{C^{s+1}(u) : u \in V(G)\} = \{C^{s+1}(v) : v \in V(H)\}. \quad (12)$$

Note that Equality (12) does not require that each color occurs in G and H the same number of times.

Proof. If G and H have different number of vertices, these graphs are distinguished by a sentence of quantifier depth 1 in $\text{FO}_{\#}^2$. If Equality (12) is not true, G and H can be distinguished by a $\text{FO}_{\#}^2$ -sentence of quantifier depth $s + 2$ by Lemma 4.1.

Conversely, assume that G and H have equal number of vertices and Equality (12) is true. Then $\{C^s(u) : u \in V(G)\} = \{C^s(v) : v \in V(H)\}$ because differently colored vertices cannot get equal colors later. Since $s = \text{Stab}(G)$, for every $c, c' \in \{C^s(u) : u \in V(G)\}$ any vertex $u \in V(G)$ with $C^s(u) = c$ has the same number $m_{c,c'}$ of neighbors w such that $C^s(w) = c'$. The matrix $M = (m_{c,c'})$ is called the *degree refinement matrix* of G . Note that the same property, with the same matrix M , holds true also for H for else (12) would be false (this implies, in particular, that $\text{Stab}(H) \leq s$).

Now, let us show that

$$\{\{C^s(u) : u \in V(G)\}\} = \{\{C^s(v) : v \in V(H)\}\}, \quad (13)$$

which means that G and H contain the same number of vertices of each C^s -color c . Regarded as an adjacency matrix, the matrix M determines directed multigraphs on the vertex sets $\{C^s(u) : u \in V(G)\}$ and $\{C^s(v) : v \in V(H)\}$. Since at least one of G and H is connected, the directed multigraphs determined by M are connected too. Assume for a while that there is a color c such that $|\{u \in V(G) : C^s(u) = c\}| < |\{v \in V(H) : C^s(v) = c\}|$. Note that the inequality $|\{u \in V(G) : C^s(u) = c'\}| < |\{v \in V(H) : C^s(v) = c'\}|$ is true also for any adjacent color c' (such that $m_{c,c'} \neq 0$). It follows by connectedness that $|V(G)| < |V(H)|$. This contradiction proves (13).

Using (13), we can show that Duplicator has a winning strategy in the counting game on G and H for any number of rounds. Assume that the pebble q occupies the vertices u in G and v in H such that $C^s(u) = C^s(v)$; note that the condition (13) allows Duplicator to ensure this in the first round. Assume that Spoiler now

plays with the other pebble p and specifies a set A . Duplicator is able to respond with a set B such that, for each C^s -color c , the number of vertices in B colored in c is the same as the number of such vertices in A . Moreover, based on the fact that G and H have the same degree refinement matrix M , Duplicator can ensure that the same is true even if only the neighbors or only the non-neighbors of u and v are considered. This allows Duplicator not to lose this round and also to ensure that the vertices under p in G and H will have the same C^s -color. ■

Lemma 4.4. *If G and H are distinguishable in $\text{FO}_{\#}^2$, then $D(G, H) \leq \text{Stab}(G) + 2$.*

Proof. Note that $D(G, H) = D(\overline{G}, \overline{H})$, where \overline{G} and \overline{H} denote the complements of G and H respectively. Considering the complements if necessary, we therefore can suppose that G is connected. Thus, we are in the conditions of Lemma 4.3. Let $s = \text{Stab}(G)$. Lemma 4.3 implies that G and H can be distinguished by specifying a C^{s+1} -color c occurring in only one of these graphs. By Lemma 4.1, the existence of a vertex colored in s can be expressed by a statement of quantifier depth $s + 2$. ■

Theorem 4.5. *The problem of deciding if two given connected graphs have a common cover is P-complete under AC^0 -reductions.*

Proof. The problem is in P due to Angluin's algorithm; see Section 2.4. The hardness for P follows by reduction from the $\text{FO}_{\#}^2$ -equivalence problem, which is to decide whether two given graphs G and H are distinguishable or equivalent in $\text{FO}_{\#}^2$. Note that this problem is equivalent to its restriction to connected graphs. Lemmas 2.8 and 4.3 readily imply that connected G and H with the same number of vertices are $\text{FO}_{\#}^2$ -equivalent if and only if they have a common covering graph. The P-completeness of the $\text{FO}_{\#}^2$ -equivalence problem is established by Grohe [20].² ■

The following result shows that the $\text{FO}_{\#}^2$ -equivalence type of a graph with n vertices is definable with quantifier depth at most $n + 1$ and that this bound is asymptotically tight.

Theorem 4.6.

1. *If n -vertex graphs G and H are distinguishable in $\text{FO}_{\#}^2$, then $D(G, H) \leq n + 1$.*
2. *For each n , there are n -vertex graphs G and H distinguishable in $\text{FO}_{\#}^2$ such that $D(G, H) > n - 8\sqrt{n}$.*

²The P-completeness of the $\text{FO}_{\#}^2$ -equivalence is stated in [20] for directed graphs with noting that the proof works also for (undirected) vertex-colored graphs. The $\text{FO}_{\#}^2$ -equivalence problem for colored graphs easily reduces to the version of this problem for uncolored graphs. Given a colored graph G with n vertices, attach $n + i$ new vertices to each vertex of G colored in the i -th color, remove all colors, and denote the resulting graph by G' . Then G and H are equivalent in $\text{FO}_{\#}^2$ if and only if G' and H' are equivalent in this logic. This easily follows from the characterization of the $\text{FO}_{\#}^2$ -equivalence by the bijection game suggested by Hella [21].

The proof of Theorem 4.6 takes the rest of this section. Part 1 follows immediately from Lemma 4.4. To prove Part 2, we consider the same graphs $G_{s,t}$ and $H_{s,t}$ as in Section 3. By Lemma 4.2, it suffices to show that Spoiler has a winning strategy in the 2-pebble counting game on these graphs but Duplicator is able to resist longer than $n - 8\sqrt{n}$ rounds. The former fact (Spoiler wins) follows from the proof of Lemma 3.3.1; however, we give a much shorter and simpler argument for it. More effort is needed to prove the latter fact (Duplicator resists for long); the game analysis is now harder because Spoiler has more freedom than in the bisimulation version.

We first design a winning strategy for Spoiler in the 2-pebble counting game on $G_{s,t}$ and $H_{s,t}$, showing that these graphs are distinguishable in $\text{FO}_{\#}^2$. Let us use the notation and the notions introduced in Section 3 for analysis of the bisimulation version of this game.

In the first round, let Spoiler pebble a vertex a_1 in $G_{s,t}$ at the maximum distance from the vertex u . That is, $\ell(a_1) = l + 5$ where l is defined by (3). Whatever Duplicator's response b_1 in $H_{s,t}$ is, $\ell(b_1) \leq l + 2$. In the subsequent rounds, Spoiler pebbles, one by one, adjacent vertices along a shortest path from b_1 to v . When Spoiler reaches the vertex v of degree 1, Duplicator arrives at a vertex of degree at least 2 in $G_{s,t}$. Spoiler wins in the next round.

In order to prove the lower bound for $D(G, H)$, we design a strategy for Duplicator allowing her to stay alive for a long time. This strategy consists in ensuring that

- a_1 and b_1 are of the same type and, furthermore, a_i and b_i are of the same type as long as

$$\min\{\ell(a_{i-1}), \ell(b_{i-1})\} > 0 \text{ or } \ell(a_{i-1}) = \ell(b_{i-1}) = 0. \quad (14)$$

The condition (14) ensures that, if a_{i-1} and b_{i-1} are of the same type, then they have equally many neighbors and non-neighbors of each type and, hence, Duplicator not only is able not to lose the next i -th round but even to secure a_i and b_i are of the same type. To keep (14) true as long as possible, Duplicator tries in each round to fulfil at least one of the following three conditions

$$\ell(a_i) = \ell(b_i) \quad (15)$$

$$\min\{\ell(a_i), \ell(b_i)\} \geq l - 1 \quad (16)$$

$$\min\{\ell(a_i), \ell(b_i)\} \geq \min\{\ell(a_{i-1}), \ell(b_{i-1})\} - 1 \quad (17)$$

Let us use an inductive argument to show that such a strategy does exist.

Lemma 4.7.

1. *In the first round of the Immerman-Lander game on $G_{s,t}$ and $H_{s,t}$, Duplicator is able to ensure pebbling vertices a_1 and b_1 of the same type satisfying at least one of the conditions (15) and (16) for $i = 1$.*

2. If a_{i-1} and b_{i-1} are of the same type and $\ell(a_{i-1}) = \ell(b_{i-1})$, then in the next round Duplicator is able to ensure pebbling vertices a_i and b_i of the same type satisfying at least one of the conditions (15) and (16).
3. If a_{i-1} and b_{i-1} are of the same type, $\ell(a_{i-1}) \neq \ell(b_{i-1})$, and $\min\{\ell(a_{i-1}), \ell(b_{i-1})\} > 1$, then in the next round Duplicator is able to ensure pebbling vertices a_i and b_i of the same type satisfying at least one of the three conditions (15)–(17).

Proof. 1. Let $\phi : V(G_{s,t}) \rightarrow V(H_{s,t})$ be a bijection that preserves vertex types and is a partial isomorphism from $G_{s,t}$ to $H_{s,t}$ up to level $l - 1$. Note that, for any $a \in V(G_{s,t})$,

$$\ell(a) = \ell(\phi(a)) \text{ or } \min\{\ell(a), \ell(\phi(a))\} \geq l. \quad (18)$$

Duplicator mirrors Spoiler's move according to ϕ , that is, after Spoiler specifies a set A , Duplicator responds with $B = \phi(A)$ if $A \subseteq V(G_{s,t})$ or $B = \phi^{-1}(A)$ if $A \subseteq V(H_{s,t})$. After Spoiler pebbles $b \in B$, Duplicator pebbles the vertex $a \in A$ such that $a = \phi^{-1}(b)$ or $a = \phi(b)$ respectively.

2. If a_{i-1} and b_{i-1} are pebbled in the preceding round, Duplicator has to modify ϕ to ϕ' so that

$$\phi'(a_{i-1}) = b_{i-1}. \quad (19)$$

Another condition to obey is

$$\phi'(N(a_{i-1})) = N(b_{i-1}), \quad (20)$$

which is possible because the assumption $\ell(a_{i-1}) = \ell(b_{i-1})$ implies that a_{i-1} and b_{i-1} are of the same degree. By Property A of the construction of $G_{s,t}$ and $H_{s,t}$, the map ϕ' can be defined on $N(a_{i-1})$ so that it respects the vertex types. The equality (20) ensures the condition

$$a_i \in N(a_{i-1}) \iff b_i \in N(b_{i-1}),$$

which is necessary for Duplicator's survival starting from the second round. If $\ell(a_{i-1}) = \ell(b_{i-1}) \leq l - 1$, then ϕ' can still be chosen to be a partial isomorphism up to level $l - 1$, and this case is much similar to Part 1. It is also possible that $\ell(a_{i-1}) = \ell(b_{i-1})$ is equal to l or to $l + 1$, for example, if $a_{i-1} = u'$ and $b_{i-1} = v'_j$. Then ϕ' can be supposed to be a partial isomorphism up to level $l - 2$ and then

$$\ell(a) = \ell(\phi'(a)) \text{ or } \min\{\ell(a), \ell(\phi'(a))\} \geq l - 1,$$

ensuring (15) or (16).

3. If $\ell(a_{i-1}) \neq \ell(b_{i-1})$, the modification of ϕ to ϕ' has to be described with more care. First of all, ϕ is modified on $\{a_{i-1}\} \cup N(a_{i-1})$ so that (19) and (20) are fulfilled. Suppose that, as a result, $\phi'(a) = b$ for $a \in \{a_{i-1}\} \cup N(a_{i-1})$ and $b \in \{b_{i-1}\} \cup N(b_{i-1})$. We are now forced to define the new image of $a' = \phi^{-1}(b)$ and the new preimage of $b' = \phi(a)$, and we do this by setting $\phi'(a') = b'$. Note that the vertex types are preserved.

The above definition of ϕ' is ambiguous only when a_{i-1} and b_{i-1} are of type \blacksquare and $|\ell(a_{i-1}) - \ell(b_{i-1})| = 2$. If in this case $a' \in N(a_{i-1})$, then $\phi'(a')$ is defined from the very beginning and should not be redefined any more. We remove this collision by supposing that $\phi(N(a_{i-1})) = N(b_{i-1})$ (if necessary, apply an automorphism of $G_{s,t}$ transposing two pairs of \blacksquare -vertices in the intermediate level between $\ell(a_{i-1})$ and $\ell(b_{i-1})$). Then ϕ' coincides with ϕ on $N(a_{i-1})$, and there is no need to modify ϕ further.

Assume that the i -th round has been played. If $\phi'(a_i) = \phi(a_i)$, then $a = a_i$ satisfies (18), which implies (15) or (16). If $\phi'(a_i) \neq \phi(a_i)$, then we have (17) because ϕ' differs from ϕ only on the levels neighboring with $\ell(a_{i-1})$ and $\ell(b_{i-1})$. \blacksquare

To complete the proof of Theorem 4.6, fix a strategy for Duplicator as in Lemma 4.7 and an arbitrary winning strategy for Spoiler. Assume that Spoiler wins in the $(r + 1)$ -th round. Note that

$$\ell(a_r) \neq \ell(b_r) \text{ and } \min\{\ell(a_r), \ell(b_r)\} = 0$$

because otherwise we would get a contradiction with Lemma 4.7. Let k be the smallest number such that $\ell(a_i) \neq \ell(b_i)$ for all $k \leq i \leq r$. Note that either $k = 1$ or $\ell(a_{k-1}) = \ell(b_{k-1})$. Parts 1 and 2 of Lemma 4.7 imply that $\min\{\ell(a_k), \ell(b_k)\} \geq l - 1$. Consider the largest index $m \geq k$ for which $\min\{\ell(a_m), \ell(b_m)\} \geq l - 1$. By Lemma 4.7.3, Spoiler needs no less than $l - 1$ rounds to decrease $\min\{\ell(a_i), \ell(b_i)\}$ from $l - 1$ to 0. Thus, $r \geq m + l - 1 \geq l$. Like in the proof of Theorem 3.2, we take $G = G_{2t+1,t}$ and $H = H_{2t+1,t}$, adding new dummy vertices if necessary, and conclude that $D(G, H) \geq l > n - 8\sqrt{n}$.

5 Comments and questions

1. Let G and H be connected n -vertex graphs with diameters at most D . Boldi and Vigna [9] notice that the isomorphism of the universal covers $U_x(G)$ and $U_y(H)$ is implied by the isomorphism of their truncations $U_x^{n+D}(G)$ and $U_y^{n+D}(H)$. Since $D \leq n - 1$, the upper bound of $n + D$ for the distinguishing truncation depth is more advantageous than $2n - 1$. Our result shows that the bound of $n + D$ is, in general, also asymptotically tight. If $D = (1 - o(1))n$, then a lower bound of $(1 - o(1))(n + D) = (2 - o(1))n$ is given directly by Theorem 3.2. If $D = o(n)$ (in particular, if $D = O(1)$), we can obtain a lower bound of $(1 - o(1))(n + D) = (1 - o(1))n$ by slightly modifying our construction of graphs $G_{s,t}$ and $H_{s,t}$. In order to decrease the diameter, we appropriately choose a set T of vertex types, add a new vertex to each of the graphs and connect it to all vertices whose type is in T .

Is the upper bound of $n + D$ tight if $D = cn$ for a constant $c \in (0, 1)$? For example, if $D = \frac{1}{2}n$, is then the upper bound of $\frac{3}{2}n$ tight up to $o(n)$?

2. Define $S(n)$ to be the maximum $\text{Stab}(G)$ over all graphs G with n vertices. Combining Theorem 3.2 and Lemma 2.6, we conclude that

$$n - 8\sqrt{2}\sqrt{n} < S(n) < n.$$

Is it true that $S(n) = n - O(1)$? Alternatively, does there exist a function $f(n)$ going to the infinity as n increases such that $S(n) < n - f(n)$?

3. Similarly to $\text{FO}_{\#}^2$, let $\text{FO}_{\#}^k$ denote the k -variable logic with counting quantifiers. The equivalence class of a graph G in $\text{FO}_{\#}^k$ consists of all graphs indistinguishable from G in this logic. There is a sentence Φ in $\text{FO}_{\#}^k$ that defines this class in the sense that Φ is true on G (hence on all $\text{FO}_{\#}^k$ -equivalent graphs) but false on every non- $\text{FO}_{\#}^k$ -equivalent graph. Let $D^k(G)$ denote the minimum quantifier depth of such a sentence. Define the function $D^k(n)$ as the maximum $D^k(G)$ over n -vertex graphs. Theorem 4.6 readily implies that $D^2(n) = (1 - o(1))n$.

If $k \geq 3$, it is known [13, 35] that $D^k(n) \leq n^{k-1}$ for $n \geq 2$. How tight is this bound? A linear lower bound is shown by Fürer [17].

References

- [1] A. Amit and N. Linial. Random lifts of graphs: Edge expansion. *Combinatorics, Probability & Computing*, 15(3):317–332, 2006.
- [2] D. Angluin. Local and global properties in networks of processors. In *Proc. of the 12th Annual ACM Symposium on Theory of Computing*, pages 82–93. ACM, 1980.
- [3] A. Atserias and E. N. Maneva. Sherali-Adams relaxations and indistinguishability in counting logics. *SIAM J. Comput.*, 42(1):112–137, 2013.
- [4] L. Babai, P. Erdős, and S. M. Selkow. Random graph isomorphism. *SIAM J. Comput.*, 9(3):628–635, 1980.
- [5] C. Berkholz, P. Bonsma, and M. Grohe. Tight lower and upper bounds for the complexity of canonical colour refinement. In *Algorithms — ESA 2013 — 21st Annual European Symposium. Proceedings*, volume 8125 of *Lecture Notes in Computer Science*, pages 145–156. Springer, 2013.
- [6] N. Biggs. *Algebraic graph theory*. Cambridge: Cambridge University Press, 2nd ed., 1994.
- [7] Y. Bilu and N. Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.
- [8] H. L. Bodlaender. The classification of coverings of processor networks. *J. Parallel Distrib. Comput.*, 6(1):166–182, 1989.
- [9] P. Boldi and S. Vigna. Fibrations of graphs. *Discrete Mathematics*, 243(1–3):21–66, 2002.
- [10] A. Cardon and M. Crochemore. Partitioning a Graph in $O(|A| \log_2 |V|)$. *Theor. Comput. Sci.* 19:85–98, 1982.

- [11] B. Courcelle and Y. Métivier. Coverings and minors: Application to local computations in graphs. *Eur. J. Comb.*, 15(2):127–138, 1994.
- [12] D. M. Cvetković, M. Doob, and H. Sachs. *Spectra of graphs. Theory and applications*. Leipzig: J. A. Barth Verlag, 3rd rev. ed., 1995.
- [13] A. Dawar, S. Lindell, S. Weinstein, Infinitary logic and inductive definability over finite structures. *Information and Computation* 119:160–175, 1995.
- [14] D. Dereniowski, A. Kosowski, and D. Pajak. Distinguishing views in symmetric networks: A tight lower bound. E-print: <http://arxiv.org/abs/1407.2511> (2014).
- [15] J. Fiala and J. Kratochvíl. Locally constrained graph homomorphisms — structure, complexity, and applications. *Computer Science Review*, 2(2):97–111, 2008.
- [16] M. J. Fischer, N. A. Lynch, and M. Merritt. Easy impossibility proofs for distributed consensus problems. *Distributed Computing*, 1(1):26–39, 1986.
- [17] M. Fürer. Weisfeiler-Lehman refinement requires at least a linear number of iterations. In: *Proc. of the 28th Int. Colloquium on Automata, Languages, and Programming*. Lecture Notes in Computer Science, Vol. 2076, 322–333, 2001.
- [18] C. Godsil and G. Royle. *Algebraic graph theory*. New York, NY: Springer, 2001.
- [19] E. Grädel and M. Otto. On logics with two variables. *Theor. Comput. Sci.* 224(1–2):73–113, 1999.
- [20] M. Grohe. Equivalence in finite-variable logics is complete for polynomial time. *Combinatorica*, 19(4):507–532, 1999.
- [21] L. Hella. Logical hierarchies in PTIME. *Inf. Comput.*, 129(1):1–19, 1996.
- [22] J. M. Hendrickx. Views in a graph: To which depth must equality be checked? *IEEE Transactions on Parallel and Distributed Systems*, 25(7):1907–1912, 2014.
- [23] J. Hopcroft. An $n \log n$ algorithm for minimizing states in a finite automaton. In *Theory of machines and computations*, page 189–196. Academic Press, New York, 1971.
- [24] N. Immerman and E. Lander. Describing graphs: A first-order approach to graph canonization. In *Complexity Theory Retrospective*, pages 59–81. Springer-Verlag, 1990.
- [25] P. J. Kelly. A congruence theorem for trees. *Pacific Journal of Mathematics*, 7:961–968, 1957.

- [26] E. Kranakis. Symmetry and computability in anonymous networks: A brief survey. In *Structure, Information and Communication Complexity: 3rd Colloquium. Proceedings*, pages 46–58. McGill-Queen’s Press, 1997.
- [27] F. T. Leighton. Finite common coverings of graphs. *J. Comb. Theory, Ser. B* 33(3):231–238, 1982.
- [28] I. Litovsky, Y. Métivier, and E. Sopena. Graph relabelling systems and distributed algorithms. In *Handbook of Graph Grammars and Computing by Graph Transformation. Volume 3: Concurrency, Parallelism, and Distribution*, pages 1–56. World Scientific, Singapore, 1999.
- [29] W. S. Massey. *Algebraic topology: An introduction.*, volume 56 of *Graduate Texts in Mathematics*. Springer, 5th ed., 1981.
- [30] E. F. Moore. Gedanken-experiments on sequential machines. In *Automata Studies*, volume 34 of *Annals of Mathematical Studies*, page 129–153. Princeton University Press, Princeton, N.J., 1956.
- [31] S. Negami. Graphs which have no finite planar covering. *Bull. Inst. Math., Acad. Sin.*, 16(4):377–384, 1988.
- [32] J. Nešetřil. A congruence theorem for asymmetric trees. *Pacific Journal of Mathematics*, 37:771–778, 1971.
- [33] N. Norris. Universal covers of graphs: Isomorphism to depth $n - 1$ implies isomorphism to all depths. *Discrete Applied Mathematics*, 56(1):61–74, 1995.
- [34] M. Otto. Bisimulation and coverings for graphs and hypergraphs. In *Logic and Its Applications 2013, Proceedings*, volume 7750 of *Lecture Notes in Computer Science*, pages 5–16. Springer, 2013.
- [35] O. Pikhurko and O. Verbitsky. Logical complexity of graphs: a survey. In M. Grohe and J. Makowsky, editors, *Model theoretic methods in finite combinatorics*, volume 558 of *Contemporary Mathematics*, pages 129–179. American Mathematical Society (AMS), Providence, RI, 2011.
- [36] M. V. Ramana, E. R. Scheinerman, and D. Ullman. Fractional isomorphism of graphs. *Discrete Mathematics*, 132(1-3):247–265, 1994.
- [37] K. Reidemeister. *Einführung in die kombinatorische Topologie*. Vieweg & Sohn, Braunschweig, 1932.
- [38] H. Sachs. Simultane Überlagerung gegebener Graphen. *Publ. Math. Inst. Hung. Acad. Sci., Ser. A*, 9:415–427, 1965.
- [39] J. R. Stallings. Topology of finite graphs. *Inventiones Mathematicae*, 71:551–565, 1983.

- [40] J. Stillwell. *Classical topology and combinatorial group theory*. New York: Springer, 2nd ed., 1993.
- [41] S. Tani. Compression of view on anonymous networks — folded view. *IEEE Trans. Parallel Distrib. Syst.* 23(2):255–262, 2012.
- [42] M. Yamashita and T. Kameda. Computing on an anonymous network. In *Proc. of the 7th Annual ACM Symposium on Principles of Distributed Computing*, pages 117–130. ACM, 1988.